Geophysical Journal International

Geophys. J. Int. (2010) **182,** 1174–1188 doi: 10.1111/j.1365-246X.2010.04686.x

Non-linear model error and resolution properties from two-dimensional single and joint inversions of direct current resistivity and radiomagnetotelluric data

Thomas Kalscheuer,[1] María de los Ángeles García Juanatey,[2] Naser Meqbel[3]* and Laust B. Pedersen[2]

[1] *Institute of Geophysics, ETH Zürich, Sonneggstr.* 5, 8092 *Zürich, Switzerland. E-mail: Thomas.Kalscheuer@aug.ig.erdw.ethz.ch*
[2] *Department of Earth Sciences, Uppsala University, Villavägen* 16, 75236 *Uppsala, Sweden*
[3] *College of Oceanic and Atmospheric Sciences, Oregon State University,* 104 *COAS Administration Building, Corvallis,* OR 97331-5503, *USA*

Accepted 2010 May 30. Received 2010 March 28; in original form 2008 November 11

SUMMARY

For the first time, a comparative analysis of the resolution and variance properties of 2-D models of electrical resistivity derived from single and joint inversions of dc resistivity (DCR) and radiomagnetotelluric (RMT) measurements is presented.

DCR and RMT data are inverted with a smoothness-constrained 2-D scheme. Model resolution, model variance and data resolution analyses are performed both with a classical linearized scheme that employs the smoothness-constrained generalized inverse and a non-linear truncated singular value decomposition (TSVD). In the latter method, the model regularization used in the inversion is avoided and non-linear semi-axes give an approximate description of the non-linear confidence surface in the directions of the model eigenvectors. Hence, this method analyses the constraints that can be provided by the data. Model error estimates are checked against improved and independent estimates of model variability from most-squares inversions.

For single and joint inverse models of synthetic data sets, the smoothness-constrained scheme suggests relatively small model errors (typically up to 30 to 40 per cent) and resolving kernels that are spread over several cells in the vicinity of the investigated cell. Linearized smoothness-constrained errors are in good agreement with the corresponding most-squares errors. The variability of the RMT model as estimated from non-linear semi-axes is confirmed by TSVD-based most-squares inversions for most model cells within the depth range of investigation. In contrast to this, most-squares errors of the DCR model are consistently larger than errors estimated from non-linear semi-axes except for the smallest truncation levels.

The model analyses confirm previous studies that DCR data can constrain resistive and conductive structures equally well while RMT data provide superior constraints for conductive structures. The joint inversion can improve error and resolution of structures which are within the depth ranges of exploration of both methods. In such parts of the model which are outside the depth range of exploration for one method, error and resolution of the joint inverse model are close to those of the best single inversion result subject to an appropriate weighting of the different data sets.

Key words: Instability analysis; Inverse theory; Numerical approximations and analysis; Electromagnetic theory; Magnetotelluric.

1 INTRODUCTION

The solution of an inverse problem involves several steps (Jackson 1973; Parker 1977): (1) questing the existence of a solution under the simplifying assumptions made in the inverse modelling process, (2) the construction of a solution under the given assumptions, (3) assessing the stability of the solution, that is, to estimate how errors in the data and *a priori* information are reflected in errors in the model parameters and (4) to assess the degree of non-uniqueness, that is, to ask which structures in a model are required to satisfy the data given (a) that the number of data is finite, (b) that the data

*Formerly at: GeoForschungsZentrum Potsdam, Geophysical Deep Sounding, Telegrafenberg, 14473 Potsdam, Germany.

contain noise and (c) that there is some arbitrariness in the chosen model discretization or mathematical formalism.

Under the assumption of a 2-D subsurface with isotropic and frequency independent electric resistivity, Sasaki (1989), Monteiro Santos *et al.* (2007) and Candansayar & Tezkan (2008) describe the construction of smoothness-constrained joint inverse models from (R)MT and dc resistivity (DCR) data with overlapping depth ranges of investigation. In the presence of 3-D effects, it is an important question whether transverse electric (TE) mode or transverse magnetic (TM) mode radiomagnetotelluric (RMT) data should be jointly inverted with DCR data. Differences in the nature of 3-D effects might render the TE- or TM-mode data set incompatible with DCR data. Monteiro Santos *et al.* (2007) proposes to use TM-mode data in the joint inversion, as the direction of current flow in the subsurface is pre-dominantly directed along the profile in both TM-mode and DCR. In the presence of anisotropic or frequency-dependent model parameters, the combination of diffusive electromagnetic and geoelectric methods may make it necessary to allow for anisotropy or frequency dependence to avoid inconsistencies in the modelling process (Haber & Oldenburg 1997; Christensen 2000).

For linear inverse problems, measures of stability and non-uniqueness are the variance and resolution of model parameters, respectively, which are computed from the data kernel and its generalized inverse (Menke 1989). For non-linear inverse problems, a similar approach is used after linearization around a given solution (so-called linear studies). The validity of approximate linearity in a sufficiently large range around the solution needs to be tested. Therefore, alternative approaches to assess non-uniqueness in non-linear inverse problems are based on the search for (possibly extremal) models that fit the data to an appropriate misfit threshold (so-called non-linear studies).

Linear studies are based entirely on sensitivity analyses (Schwalenberg *et al.* 2002) or on both the sensitivity matrix and its generalized inverse (Alumbaugh & Newman 2000; Friedel 2003). In contrast to Alumbaugh & Newman (2000) who derive model variance and resolution for a smoothness-constrained inverse scheme, Friedel (2003) introduces an singular value decomposition (SVD)-based analysis of data resolution, model resolution and model covariance for the DCR inverse problem. In non-linear studies, different *a priori* conditions are used to obtain model solutions that are not linearly close to each other (Oldenburg & Li 1999; Schwalenberg *et al.* 2002). Non-uniqueness is then characterized through the differences of the inverse models. Muñoz & Rath (2006) utilize an SVD-based null-space projection-scheme to construct equivalent models that have the same misfit as a given inverse model. Another approach to non-linear studies is to construct extremal models that fit the data to a given misfit threshold with the edgehog method (Jackson 1973; Lines & Treitel 1985) or the most-squares method (Jackson 1976; Meju & Hutton 1992; Meju 1994, 2009). Miller & Routh (2007) establish a theoretical relationship between the depth-of-investigation (DOI) index derived by Oldenburg & Li (1999) from non-linear studies and the model resolution matrix derived from linear studies.

Kalscheuer & Pedersen (2007) estimate model error and resolution with a truncated singular value decomposition (TSVD) analysis, where inverse singular values are replaced by non-linear semi-axes to a given level of the non-linear confidence surface. A model error threshold determines the truncation level and hence the model resolution matrix. Error estimates are checked with a most-squares inversion. The TSVD analysis by Kalscheuer & Pedersen (2007) avoids the smoothness constraints used during the inverse process. Hence, it analyses to which extent the data alone can constrain the model. Ideally, for this purpose, the model resolution–covariance trade-off curve employed in the TSVD analysis would be investigated in greater detail. Due to the high computational cost involved in determining non-linear semi-axes and most-squares error estimates, investigating a large number of error levels is prohibitive.

In this paper, 2-D models of electric resistivity are constructed with a smoothness-constrained inverse scheme that is applied to the single and joint inversion of RMT and DCR data. As a novelty, we quantitatively deduce model variance, model resolution and data resolution properties of the 2-D models with two schemes, that is, a linear study that employs the generalized inverse from a smoothness-constrained inversion and the non-linear TSVD scheme by Kalscheuer & Pedersen (2007). In both schemes, the estimated model errors are verified with iterative most-squares inversions. Hence, there are two further improvements. First, the smoothness-constrained most-squares inversion is employed to validate the widely used linearized smoothness-constrained error estimates. Second, the non-linear behaviour of the DCR and RMT inverse problems can be assessed.

2 INVERSE THEORY

The iterative least-squares solution in the $(k+1)$th iteration of a non-linear inverse problem can be expressed as (Menke 1989; Siripunvaraporn & Egbert 2000)

$$\begin{aligned}\mathbf{m}_{k+1}(\lambda) &= \left[\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_{m}^{\mathrm{T}}\mathbf{W}_{m}\right]^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\hat{\mathbf{d}}_{\mathrm{k}} + \mathbf{m}_{\mathrm{r}} \\ &= \mathbf{J}_{\mathbf{W}}^{-\mathrm{g}}\mathbf{W}_{\mathrm{d}}\hat{\mathbf{d}}_{\mathrm{k}} + \mathbf{m}_{\mathrm{r}}. \end{aligned} \tag{1}$$

Here, $\mathbf{J}_{\mathbf{W}}^{-\mathrm{g}} = [\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_{m}^{\mathrm{T}}\mathbf{W}_{m}]^{-l}\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}$ is the generalized inverse, $\mathbf{J}$ is the sensitivity matrix, $\mathbf{W}_{\mathrm{d}}$ is a data weighting matrix, $\mathbf{W}_{\mathrm{m}}$ is a model regularizing matrix, $\mathbf{m}_{\mathrm{r}}$ is a reference model and $\hat{\mathbf{d}}_k = \mathbf{d} - \mathbf{F}[\mathbf{m}_k] + \mathbf{J}(\mathbf{m}_k - \mathbf{m}_{\mathrm{r}})$ is a data difference vector with the non-linear forward response $\mathbf{F}[\mathbf{m}_k]$ of the model $\mathbf{m}_k$ of the kth iteration. Further details on the terminology used in this paper and the employed model regularizing schemes can be found in Appendix A. Details on forward computations, sensitivity computations and data weighting are presented in Appendix B.

2.1 Model resolution, model covariance and data resolution matrices

The estimated model $\mathbf{m}_{k+1}$ can be related to the unknown true model $\mathbf{m}^{\mathrm{true}}$ by assuming that the measured data $\mathbf{d}$ represent the responses of the true model and a noise term $\mathbf{n}$ and that the response of the model $\mathbf{m}_k$ is linearly close to that of the true model (Friedel 2003),

$$\begin{aligned}\mathbf{d} &= \mathbf{F}[\mathbf{m}^{\mathrm{true}}] + \mathbf{n} \\ &\approx \mathbf{F}[\mathbf{m}_k] + \mathbf{J}(\mathbf{m}^{\mathrm{true}} - \mathbf{m}_{\mathrm{k}}) + \mathbf{n}. \end{aligned} \tag{2}$$

Insertion of eq. (2) into eq. (1) yields (*cf.* Menke 1989):

$$\mathbf{m}_{k+1} \approx \mathbf{R}_{\mathrm{M}}\mathbf{m}^{\mathrm{true}} + (\mathbf{I} - \mathbf{R}_{\mathrm{M}})\,\mathbf{m}_{\mathrm{r}} + \mathbf{J}_{\mathbf{W}}^{-\mathrm{g}}\mathbf{n}_{\mathbf{W}}, \tag{3}$$

where $\mathbf{R}_{\mathrm{M}} = \mathbf{J}_{\mathbf{W}}^{-\mathrm{g}}\mathbf{J}_{\mathbf{W}}$ is the model resolution matrix and $\mathbf{n}_{\mathbf{W}} = \mathbf{W}_{\mathrm{d}}\mathbf{n}$.

The jth row of the model resolution matrix $\mathbf{R}_{\mathrm{M}}$ describes the weights with which the parameters of the true model enter into the jth parameter of the estimated model. If, for instance, the diagonal entry $R_{\mathrm{M},jj}$ is the only non-zero entry in the jth row, only the jth parameter of the true model $\mathbf{m}^{\mathrm{true}}$ and the noise term will contribute to the jth parameter of $\mathbf{m}_{k+1}$. In such a case, the jth parameter is said to be perfectly resolved. In contrast to this, if there is a large number of non-zero entries in the jth row with approximately

equal amplitudes, the jth parameter of $\mathbf{m}_{k+1}$ will be an average over a wide range of parameters of the true model and is also strongly influenced by the chosen reference model. Then, the jth parameter is said to be badly resolved. Perfect resolution of all model parameters ($\mathbf{R}_\mathrm{M} = \mathbf{I}$) can only be obtained for overdetermined least-squares problems (Menke 1989).

As explained by Kalscheuer & Pedersen (2007), the jth row of $\mathbf{R}_\mathrm{M}$ contains the discrete resolving kernel $r_{\mathrm{M}j}$ multiplied by the respective cell areas. We depict the entries of $r_{\mathrm{M}j}$ scaled by its maximum modulus.

The model covariance matrix is a measure of stability of the inverse model that describes how data errors (given by $\mathbf{W}_\mathrm{d}$) and errors in the reference model propagate into errors of $\mathbf{m}_{k+1}$ (Menke 1989). The model covariance matrix can be directly derived from eq. (3) as

$$\begin{aligned}[\mathrm{cov}\,\mathbf{m}_{k+1}] &= E\left[(\mathbf{m}_{k+1} - E[\mathbf{m}_{k+1}])(\mathbf{m}_{k+1} - E[\mathbf{m}_{k+1}])^\mathrm{T}\right] \\ &\approx (\mathbf{I} - \mathbf{R}_\mathrm{M})[\mathrm{cov}\,\mathbf{m}_r](\mathbf{I} - \mathbf{R}_\mathrm{M})^\mathrm{T} + \mathbf{J}_\mathbf{W}^{-\mathrm{g}}\mathbf{J}_\mathbf{W}^{-\mathrm{g}^\mathrm{T}},\end{aligned} \tag{4}$$

where $E[\mathbf{A}]$ denotes the expectation value of an arbitrary matrix $\mathbf{A}$. The covariances of the true model and the data are assumed to be $[\mathrm{cov}\,\mathbf{m}^\mathrm{true}] = 0$ and $[\mathrm{cov}\,\mathbf{d}] = [\mathrm{cov}\,\mathbf{n}] = (\mathbf{W}_\mathrm{d}^\mathrm{T}\mathbf{W}_\mathrm{d})^{-1}$, respectively. The covariance matrix of the reference model is $[\mathrm{cov}\,\mathbf{m}_r] = (\lambda\mathbf{W}_m^\mathrm{T}\mathbf{W}_\mathrm{m})^{-1}$. In the smoothness-constrained and Marquardt–Levenberg schemes, $\mathbf{m}_r$ is typically reckoned a fixed vector and, hence, $[\mathrm{cov}\,\mathbf{m}_r] = 0$. The standard deviation of the jth model parameter is the square root of the jth diagonal entry of eq. (4). The errors of the logarithmic cell resistivities used in the inversion translate to error factors f on actual cell resistivities corresponding to ranges $[\rho/f, f\rho]$.

The data $\mathbf{d}_{k+1}^\mathrm{pre}$ predicted by the inverse model $\mathbf{m}_{k+1}$ can be related to the measured data $\mathbf{d}$ through linearization and substitution of eq. (1):

$$\begin{aligned}\mathbf{d}_{k+1}^\mathrm{pre} &= \mathbf{F}[\mathbf{m}_{k+1}] \approx \mathbf{F}[\mathbf{m}_k] + \mathbf{J}(\mathbf{m}_{k+1} - \mathbf{m}_k) \\ &= \mathbf{R}_\mathrm{D}\mathbf{d} + (\mathbf{I} - \mathbf{R}_\mathrm{D})(\mathbf{F}[\mathbf{m}_k] - \mathbf{J}(\mathbf{m}_k - \mathbf{m}_r)).\end{aligned} \tag{5}$$

Here, $\mathbf{R}_\mathrm{D} = \mathbf{J}\mathbf{J}_\mathbf{W}^{-\mathrm{g}}\mathbf{W}_\mathrm{d}$ is the data resolution matrix that describes the fit of the predicted data $\mathbf{d}_{k+1}^\mathrm{pre}$ to the field data $\mathbf{d}$. For completely underdetermined problems, $\mathbf{R}_\mathrm{D} = \mathbf{I}$ and perfect data fit is obtained, that is, $\mathbf{d}_{k+1}^\mathrm{pre} = \mathbf{d}$ (Menke 1989). For mixed-determined and overdetermined problems, the predicted data generally contain terms related to the field data, the model of the kth iteration and the reference model.

If the response of model $\mathbf{m}_k$ is linearly close to that of the reference model, eq. (5) simplifies to $\mathbf{d}_{k+1}^\mathrm{pre} \approx \mathbf{R}_\mathrm{D}\mathbf{d} + (\mathbf{I} - \mathbf{R}_\mathrm{D})\mathbf{F}[\mathbf{m}_r]$. Hence, the responses $\mathbf{d}_{k+1}^\mathrm{pre}$ that are not resolved by the field data are determined by the responses $\mathbf{F}[\mathbf{m}_r]$ of the reference model.

For a damped TSVD inversion (*cf.* Appendix A2),

$$\mathbf{R}_\mathrm{M} = \mathbf{V}_p\tilde{\mathbf{\Lambda}}_p^{-1}\mathbf{\Lambda}_p\mathbf{V}_p^\mathrm{T}, \tag{6}$$

$$[\mathrm{cov}\,\mathbf{m}_{k+1}] = \mathbf{V}_p\tilde{\mathbf{\Lambda}}_p^{-2}\mathbf{V}_p^\mathrm{T}, \tag{7}$$

$$\mathbf{R}_\mathrm{D} \approx \mathbf{W}_\mathrm{d}^{-1}\mathbf{U}_p\mathbf{\Lambda}_p\tilde{\mathbf{\Lambda}}_p^{-1}\mathbf{U}_p^\mathrm{T}\mathbf{W}_\mathrm{d}. \tag{8}$$

2.2 Shortcomings of standard variance and resolution estimates

Eqs (3)–(5) provide a theoretical framework to compute model variance, model resolution and data resolution properties for non-linear inverse problems based on the data, the reference model, the data weighting matrix, the model regularizing matrix and the Lagrange multiplier. However, there are several possibly false assumptions and shortcomings of this linearized formulation:

(1) The reference model might represent a false presumption of the subsurface. The model regularizing matrix [e.g. the matrix $\mathbf{S}$ of smoothness constraints in Appendix (A1)] might represent a mathematical construct that solely facilitates the solution finding process and is not necessarily coupled to geoelectrical reality. Hence, it is desirable to evaluate to which degree the subsurface is constrained by the data alone.

(2) The assumption of a Gaussian probability distribution as implied by the l_2-norm measure of linearized data fit in eq. (A5) (Menke 1989) is violated by the non-linearity of the forward problem. Non-linearity is only accounted for through linearization. The study by Kalscheuer & Pedersen (2007) shows that the non-linear semi-axes of the RMT problem are smaller than the linear semi-axes (i.e. the inverse singular values) after a certain singular value number. Hence, the non-linearity of this specific inverse problem leads to models that are better constrained through the data than indicated by linear studies.

(3) The directions of the model correction steps of the early inverse iterations can be quite different from those of later iterations. In terms of model eigenvectors, this means that the effective number of model eigenvectors used in the iterative process may be much larger than that of the last iteration. Hence, model resolution may be better than indicated by the linearized estimates.

2.3 Non-linear TSVD analysis

Following Kalscheuer & Pedersen (2007), we compute model variance and resolution estimates that partly allow for the non-linearity of the underlying inverse problem. As the scheme is based on the SVD, it presents an effective means to analyse how well the model is constrained by the data.

For a linear inverse problem $\mathbf{W}_\mathrm{d}\mathbf{d} = \mathbf{W}_\mathrm{d}\mathbf{G}\mathbf{m}$ with data kernel $\mathbf{G}$, a perturbation $\Delta\mathbf{m}$ away from the set of model parameters $\mathbf{m}_\mathrm{d}^\mathrm{opt}$ in which $Q_\mathrm{d} = (\mathbf{d} - \mathbf{G}\mathbf{m})^\mathrm{T}\mathbf{W}_\mathrm{d}^\mathrm{T}\mathbf{W}_\mathrm{d}(\mathbf{d} - \mathbf{G}\mathbf{m})$ attains its minimum $Q_\mathrm{d}^\mathrm{opt} = Q_\mathrm{d}[\mathbf{m}_\mathrm{d}^\mathrm{opt}]$ results in a modified data misfit (Johansen 1977)

$$\begin{aligned}Q_\mathrm{d}[\mathbf{m}] &= Q_\mathrm{d}\left[\mathbf{m}_\mathrm{d}^\mathrm{opt}\right] + \Delta\mathbf{m}^\mathrm{T}\mathbf{V}_p\mathbf{\Lambda}_p^2\mathbf{V}_p^\mathrm{T}\Delta\mathbf{m} \\ &= Q_\mathrm{d}^\mathrm{opt} + \Delta Q.\end{aligned} \tag{9}$$

The quantity $\Delta Q = \Delta\mathbf{m}^\mathrm{T}\mathbf{V}_p\mathbf{\Lambda}_p^2\mathbf{V}_p^\mathrm{T}\Delta\mathbf{m}$ describes the linear confidence surface in parameter space as a hyperellipsoid with semi-axes $s_i^\mathrm{lin} = \sqrt{\Delta Q}/\lambda_i$ (ith singular value λ_i) and axes parallel to the model eigenvectors in $\mathbf{V}_p$ (Johansen 1977).

A simplified description of the non-linear confidence surface can be obtained by determining the lengths of the actual semi-axes $s_i^\pm$ in both positive and negative directions of an eigenvector $\mathbf{v}_i$ (Johansen 1977; Pedersen 1979; Pedersen & Rasmussen 1989). The actual shape of the non-linear confidence surface in directions other than those of the eigenvectors is neglected and approximated as a pseudo-hyperellipsoid. Fig. 1 depicts a pseudo-hyperellipsoid for the simple example of two model parameters. The actual semi-axes $s_1^\pm$ and $s_2^\pm$ correspond to the points in model space where the pseudo-hyperellispoidal surface (thin blue line) intersects the true non-linear confidence surface (thick red line) for a given ΔQ.

For the linear case with truncation level p, the confidence limits of the jth model parameter are computed as (Johansen 1977; Pedersen

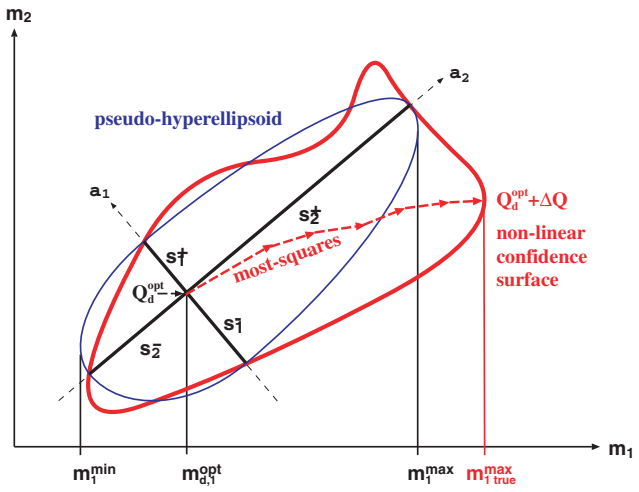

Figure 1. Illustration of the pseudo-hyperellipsoidal surface as an approximation to the true confidence surface of the non-linear problem (modified after Pedersen & Rasmussen 1989). The thick red line depicts the true confidence surface. The pseudo-hyperellipsoid is described by the thin blue line and constructed by sections of four hyperellipsoids with semi-axes (s_1^+, s_2^-), (s_1^+, s_2^+), (s_1^-, s_2^+) and (s_1^-, s_2^-). In the computation of model errors, the non-linear semi-axes $s_i^{\pm}$ for $\Delta Q = 1$ replace the linear semi-axes $s_i^{\rm lin} = \sqrt{\Delta Q}/\lambda_i$ (with singular value λ_i). The axes a_1 and a_2 point into the directions of the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively. The dashed line illustrates the iterative course of the most-squares inversion to find the true maximal value of m_1 under the constraint that the misfit equals $Q_{\rm d}^{\rm opt} + \Delta Q$.

1979; Pedersen & Rasmussen 1989)

$$\Delta m_j = \sqrt{\Delta Q} \sqrt{\sum_{i=1}^{p} v_{ji}^2/\lambda_i^2}. \qquad (10)$$

For the choice $\Delta Q = 1$ and neglecting Marquardt-type damping for the moment, eq. (10) simplifies to the standard deviation of the model parameters [square roots of diagonal entries of eq. (7)] with a confidence level of 68 per cent. The computation of non-linear confidence limits demands that the $s_i^{\pm}$ are systematically substituted for the $s_i^{\rm lin}$ in eq. (10). An error threshold $f_{\rm thresh}$ determines the truncation level p. Further details can be found in Johansen (1977) and Kalscheuer & Pedersen (2007). As the possibly damped inverse singular values in eq. (7) are replaced by the actual semi-axes, the model resolution matrix in eq. (6) is computed as $\mathbf{R}_{\rm M} = \mathbf{V}_p \mathbf{V}_p^{\rm T}$.

Only if the estimated model $\mathbf{m}^{\rm est}$ equals the optimum model $\mathbf{m}_{\rm d}^{\rm opt}$, the non-linear and linear estimates of the semi-axes have roughly equal values where the effect of non-linearity is marginal. If $\mathbf{m}^{\rm est}$ is shifted away from $\mathbf{m}_{\rm d}^{\rm opt}$, the s_i^+ and s_i^- will be systematically shifted away from $s_i^{\rm lin}$ and non-linear effects cannot be distinguished from a displacement from $\mathbf{m}_{\rm d}^{\rm opt}$ (Kalscheuer & Pedersen 2007). A series of TSVD inversion steps and model variance and resolution analyses is applied to obtain an optimum model $\mathbf{m}_{\rm d}^{\rm opt}$ that is constructed with a truncation level corresponding to the chosen variance threshold.

2.4 Most-squares inversion

Error estimates from the linearized scheme (eq. 4) and the non-linear scheme (Section 2.3) can be verified with a most-squares inversion (Jackson 1976; Meju & Hutton 1992). Here, the aim is to find the extreme values of a functional $\mathbf{m}^{\rm T}\hat{\mathbf{n}}$ under the constraint that the misfit $Q[\mathbf{m}]$ (*cf*. Appendix. A) does not exceed some threshold Q_t. If $\hat{\mathbf{n}}$ is chosen as a unit normal vector of the hyperplane $m_j = {\rm const}$, the jth model parameter will be maximized or minimized. The threshold misfit is chosen as $Q_t = Q[\mathbf{m}^{\rm opt}] + \Delta Q$ with $\Delta Q = 1$ as in Section 2.3. For the TSVD scheme, the choice $Q_t = Q_{\rm d}[\mathbf{m}_{\rm d}^{\rm opt}] + \Delta Q$ assures direct comparability with parameter error estimates computed from non-linear semi-axes (Section 2.3). The generalized inverse is constructed with the same truncation level that was used in the final non-linear TSVD analysis. For the smoothness-constrained scheme, we set $Q_t = Q_{\rm d}[\mathbf{m}^{\rm opt}] + \lambda Q_m[\mathbf{m}^{\rm opt}] + \Delta Q$ where λ is the Lagrange multiplier of the smoothness constraints that was used to obtain the model under investigation. Fixing the Lagrange multiplier λ like this assures equivalence to the linearized scheme in eq. (4). Here, the model $\mathbf{m}^{\rm opt}$ minimizes the misfit $Q[\mathbf{m}] = Q_{\rm d}[\mathbf{m}] + \lambda Q_m[\mathbf{m}]$.

For the TSVD scheme, the iterative corrections from $\mathbf{m}_{\rm d}^{\rm opt}$ to the desired level $Q_{\rm d}[\mathbf{m}_{\rm d}^{\rm opt}] + \Delta Q$ on the non-linear confidence surface are illustrated in Fig. 1 for the case where maximal resistivity is required. A formal derivation of the most-squares algorithm is presented in Appendix C.

2.5 Model construction and analysis scheme

Similar to the TSVD scheme by Kalscheuer & Pedersen (2007), there is a pre-requisite for the computation of model errors with smoothness-constrained most-squares inversions. To not mistake a displacement of the estimated model from the one with minimal misfit for a non-linear effect, the misfit $Q[\mathbf{m}] = Q_{\rm d}[\mathbf{m}] + \lambda Q_m[\mathbf{m}]$ (not the rms) is minimized during the model construction process for a given λ. This leads to the following scheme of model construction and analysis.

Step 1. In an initial trial-and-error inversion phase, an optimal Lagrange multiplier $\lambda^{\rm opt}$ for the smoothness constraints is found such that the data misfit $Q_{\rm d}$ comes close to but does not decrease below the optimal value $N \cdot ({\rm rms}^{\rm opt})^2$ determined by the employed data reweighting matrix $\mathbf{W}_{\rm rew}$ (*cf*. eq. B4). This requirement assures that the model does not fit overduely to noise.

Step 2. In every iteration k of the actual inversion, the total misfit $Q[\mathbf{m}_k] = Q_{\rm d}[\mathbf{m}_k] + \lambda^{\rm opt} Q_m[\mathbf{m}_k]$ is minimized. In addition to the smoothness constraints in $Q_m[\mathbf{m}_k]$, Marquardt-type damping is applied. An optimal Marquardt damping factor ϵ_k is determined with a simple line search. The Marquardt damping factor will typically drop by several orders of magnitude from the first to the final iteration (Lines & Treitel 1984; Rodi & Mackie 2001). The iterative process is terminated as soon as the variation of the total misfit from one iteration to the next is smaller than a fraction of the misfit deviation $\Delta Q = 1$ targeted at in the smoothness-constrained most-squares inversion. This process yields the optimum model $\mathbf{m}^{\rm opt}$.

Step 3. With $\lambda^{\rm opt}$ and the Jacobian that eventually led to $\mathbf{m}^{\rm opt}$, compute the linearized smoothness-constrained estimates of model resolution, model covariance and data resolution with eqs (3)–(5) neglecting Marquardt damping. Starting from $\mathbf{m}^{\rm opt}$ and with fixed $\lambda^{\rm opt}$, compute improved error estimates with the smoothness-constrained most-squares inversion.

Step 4. Starting from $\mathbf{m}^{\rm opt}$, perform non-linear TSVD analyses for every model parameter of interest as detailed in Section 2.3. Then starting from the corresponding $\mathbf{m}_{\rm d}^{\rm opt}$, compute TSVD-based most-squares errors with truncation levels as determined in the non-linear TSVD analyses.

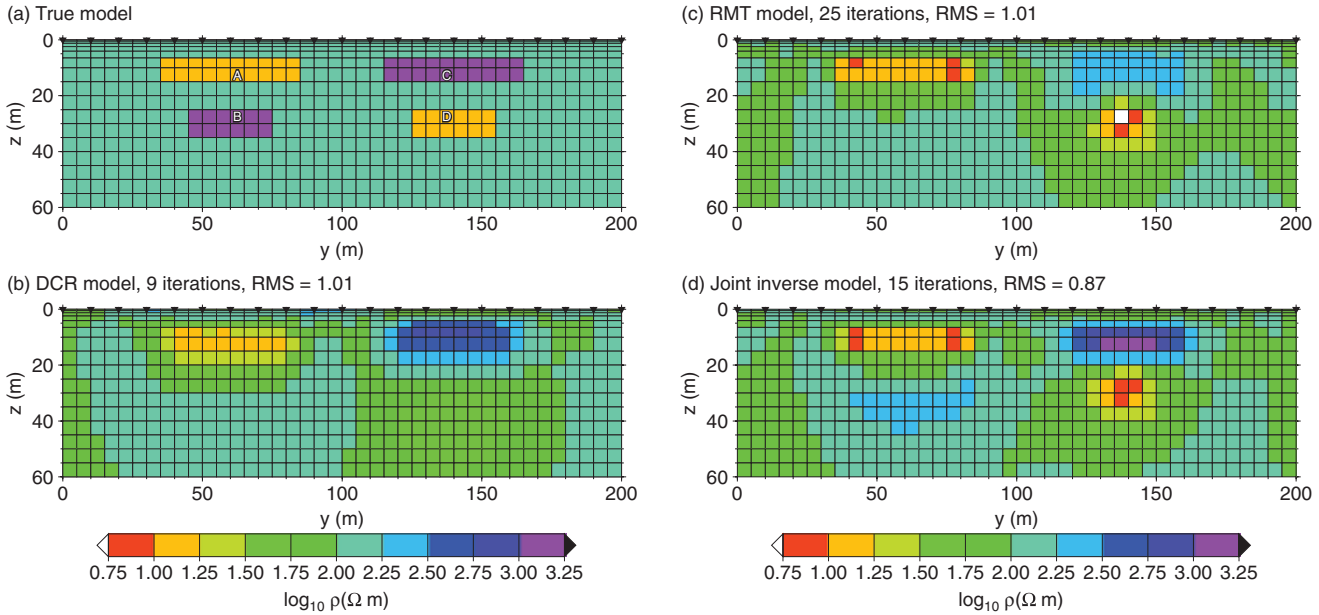

Figure 2. Models of synthetic data example from smoothness-constrained inversions: (a) true model with blocks of resistivities of 10 Ωm and 1000 Ωm in a host of 100 Ωm, (b) inverse model of DCR forward and reverse pole–dipole data, (c) inverse model of TE-mode and TM-mode data and (d) joint inverse model of TE-mode, TM-mode and DCR forward and reverse pole–dipole data. Apparently, in the joint inverse model the shallow resistive block with cell C is reproduced more accurately than in any single inversion model. To constrain the resistivity of cell D in the joint inversion similarly well as in the RMT single inversion, the weight of the DCR data and of the high-frequency RMT data at stations in the vicinity of cell D was reduced in the joint inversion.

3 A SYNTHETIC EXAMPLE

3.1 True model and synthetic data

Synthetic forward data of a 2-D model, that contains four blocks which are either conductive or resistive with respect to the host medium (Fig. 2a), were computed for the forward and reverse pole–dipole configurations and the TM and TE modes. Two blocks have resistivities of 10 Ωm, and the two other blocks have resistivities of 1000 Ωm. The host medium has a resistivity of 100 Ωm. At $y = 70$ m along the profile a conductive block is located above a resistive block and at $y = 140$ m a resistive block is located above a conductive block. RMT responses were computed at 21 receiver sites for 13 frequencies ranging from 4–250 kHz giving a total of 1092 RMT data points. Gaussian white noise corresponding to 2 per cent of the modulus of the computed impedances was added to the forward data of both polarizations. The pole–dipole array consists of 21 electrodes with a spacing of 10 m. Four expansion factors (1, 2, 4 and 6) on the basic potential electrode distance of 10 m and level values of $n = 1, \ldots, 7$ for a fixed potential electrode distance were used resulting in a total of 458 DCR data points. Gaussian white noise corresponding to 2 per cent of the apparent resistivities was added to the forward data.

Table 1. Results of model variance and resolution analyses with the smoothness-constrained scheme.

Data		Cell A	Cell B	Cell C	Cell D
DCR	Resolving kernel	Fig. 3(a)	Fig. –	Fig. 4(a)	Fig. 5(a)
	factor $f_{[\mathrm{cov\,d}]}$	1.04	1.04	1.05	1.05
	factor f_{tot}	1.17	1.25	1.20	1.25
	factor (f^-_{MSQ}/f^+_{MSQ})	1.16/1.17	1.24/1.24	1.20/1.20	1.25/1.25
TE & TM	Resolving kernel	Fig. 3(b)	Fig. –	Fig. 4(b)	Fig. 5(b)
	factor $f_{[\mathrm{cov\,d}]}$	1.06	1.07	1.06	1.07
	factor f_{tot}	1.18	1.31	1.26	1.18
	factor (f^-_{MSQ}/f^+_{MSQ})	1.16/1.16	1.29/1.28	1.25/1.26	1.35/1.34
TE & TM & DCR	Resolving kernel	Fig. 3(c)	Fig.–	Fig. 4(c)	Fig. 5(c)
	factor $f_{[\mathrm{cov\,d}]}$	1.05	1.05	1.05	1.05
	factor f_{tot}	1.14	1.22	1.18	1.16
	factor (f^-_{MSQ}/f^+_{MSQ})	1.13/1.13	1.22/1.22	1.15/1.15	1.26/1.26

Notes: The linearized error factors $f_{[\mathrm{cov\,d}]}$ are computed from last term of eq. (4). The factors f_{tot} are computed from both terms of eq. (4). The most-squares errors (f^-_{MSQ}/f^+_{MSQ}) are the non-linear equivalents of f_{tot} for decreasing (−) and increasing (+) resistivity. The figure numbers of the corresponding resolving kernels are listed. In most cases, the linearized errors f_{tot} are in good agreement with the most-squares error estimates (f^-_{MSQ}/f^+_{MSQ}). In all cases, the most-squares errors of the joint inversion are comparable to or smaller than the errors obtained from the single inversions.

3.2 Single and joint inverse models

Single and joint inversions of the synthetic data were performed with the 2-D smoothness-constrained inverse scheme. The errors were assumed to correspond to 2 per cent of the modulus of the impedances and 2 per cent of the DCR-apparent resistivities. The starting model was a homogeneous half-space of 300 Ωm. Unless stated otherwise, all inversions were run with the same weight on horizontal and vertical smoothing, in this case $\alpha_y = \alpha_z = 1$ (*cf.* Appendix A1).

For the inversion of the DCR data, a Lagrange multiplier $\lambda^{\text{opt}}_{\text{DCR}} = 49$ for the smoothness constraints proved to be optimal. After only four iterations, the inversion of the DCR data reached an acceptable data fit of rms = 1.06. However, five additional iterations were required to ascertain that the total misfit Q is reduced to its minimum such that the smoothness-constrained most-squares inversion could be applied. This final inverse model is shown in Fig. 2(b) and has an rms error of 1.00. The lateral extents of the shallow blocks are reconstructed correctly; the vertical bounds are, however, smeared out. As a consequence of the limited length of the electrode array, the deep blocks are not constrained by the data. A slight increase of the weight of horizontal smoothing results in a model (not shown) that has no evidence of the deep resistor, while fitting the data to a similar rms of 1.02.

For the inversion of the RMT data, the Lagrange multiplier of the smoothness constraints was chosen as $\lambda^{\text{opt}}_{\text{RMT}} = 28$. The inverse model of the twenty-fifth iteration (Fig. 2c) fits the data to rms = 1.01. Both conductive blocks are better resolved than in the DCR model. Below the near-surface conductor, the resolving power of the RMT data is severely diminished due to the strong damping of the electromagnetic field in the conductor. The resistive block close to the surface is poorly resolved as inductively coupled electromagnetic methods are not very sensitive to resistors (Gómez-Trevino & Edwards 1983; Raiche *et al.* 1985; Hohmann & Raiche 1987).

A joint inversion with equal weight on the two data sets yielded a model (not shown) that has better defined shallow blocks than any single inversion model. However, the upper edge of the deep conductor is not as well defined as in the inverse model of the RMT data (Fig. 2c). Hence, we decreased the weight of both the high-frequency RMT data at receiver sites above the deep conductor and the DCR data during the inversion. According to eq. (B4), the chosen weighting leads to a target rms value of 0.86. The final model of the joint inversion (Fig. 2d) has an rms misfit of 0.87 and was obtained after 15 iterations with a Lagrange multiplier of $\lambda^{\text{opt}}_{\text{joint}} = 48$. The shallow conductive and resistive blocks are better defined than in the single inversion models (Figs 2b and c). The deep conductor is reproduced with a similar accuracy as in the single inversion of RMT data (Fig. 2c).

The optimal values of the Marquardt damping factor dropped from between 160 (joint inversion) and 250 (single inversions) in the first iterations to about 10^{-2} in the final iterations.

3.3 Model error and resolution analyses

A set of four model parameters, one for each block structure, was examined with the linearized smoothness-constrained scheme (Section 2.1) and the non-linear TSVD scheme (Section 2.3). The corresponding cells are marked with labels A–D in Fig. 2(a). Cell A ($y = 62$ m, $z = 12$ m) is located in the lower part of the shallow conductive block, cell B ($y = 62$ m, $z = 37$ m) is located in the upper part of the deep resistor, cell C ($y = 137$ m, $z = 12$ m) is located in the lower part of the shallow resistor and cell D ($y = 137$ m, $z = 27$ m) is located in the upper part of the deep conductor.

For the smoothness-constrained scheme, resolution matrices and model errors were computed as explained in Sections 2.1 and 2.5. As the Marquardt damping factors were very small in the last iterations of the inversions, they were neglected during the analyses. The linearized error factors, most-squares errors and figure numbers of the corresponding resolving kernels are summarized in Table 1. The linearized errors $f_{[\text{cov d}]}$ were computed from the second term of eq. (4). However, the most-squares errors (f^-_{MSQ}/f^+_{MSQ}) are the non-linear equivalents of the errors f_{tot} computed from both terms in eq. (4) (*cf*. Appendix C). For all cells considered, the linearized standard deviations $f_{[\text{cov d}]}$ are rather small in an interval between 4.0 and 7.4 per cent. The total linearized errors f_{tot} fall into the range of 14–31 per cent and are confirmed by the most-squares errors (f^-_{MSQ}/f^+_{MSQ}) with the exemption of cell D where (f^-_{MSQ}/f^+_{MSQ}) are slightly larger for the RMT and joint inverse models. As the effect of non-linearity on model errors from TSVD analyses can be quite severe (Kalscheuer & Pedersen 2007), one might expect a

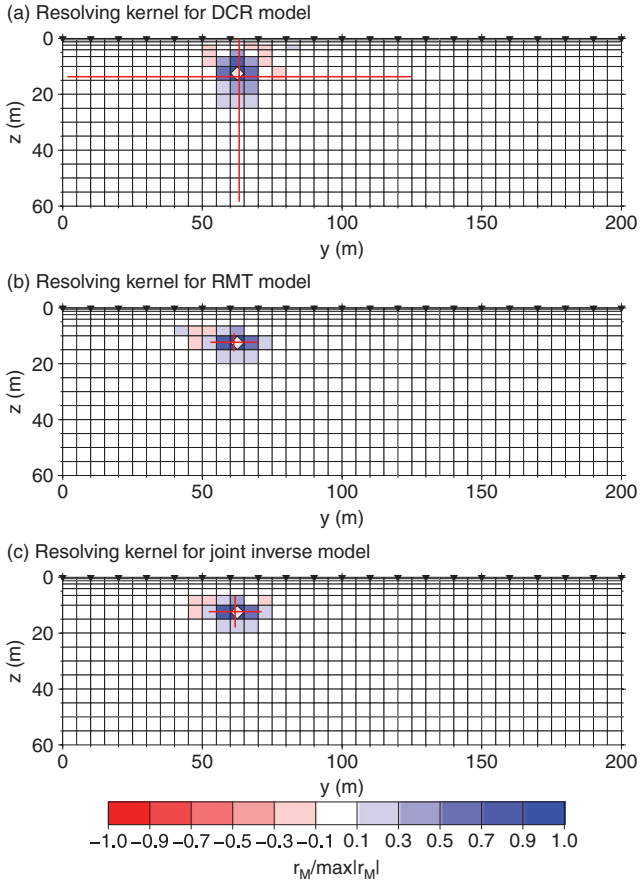

Figure 3. Resolving kernels derived from the smoothness-constrained scheme for cell A in the shallow conductor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. A white diamond marks the considered cell. The red lines depict the estimated horizontal and vertical resolution lengths; their crossing point is the estimated centre of resolution (*cf.* Appendix B in Kalscheuer & Pedersen 2007). The corresponding error estimates are given in Table 1. The RMT data offer better resolving power for the conductive structure than the DCR data. As compared to the RMT model, the resolving kernel for the joint inverse model has a marginally larger spread.

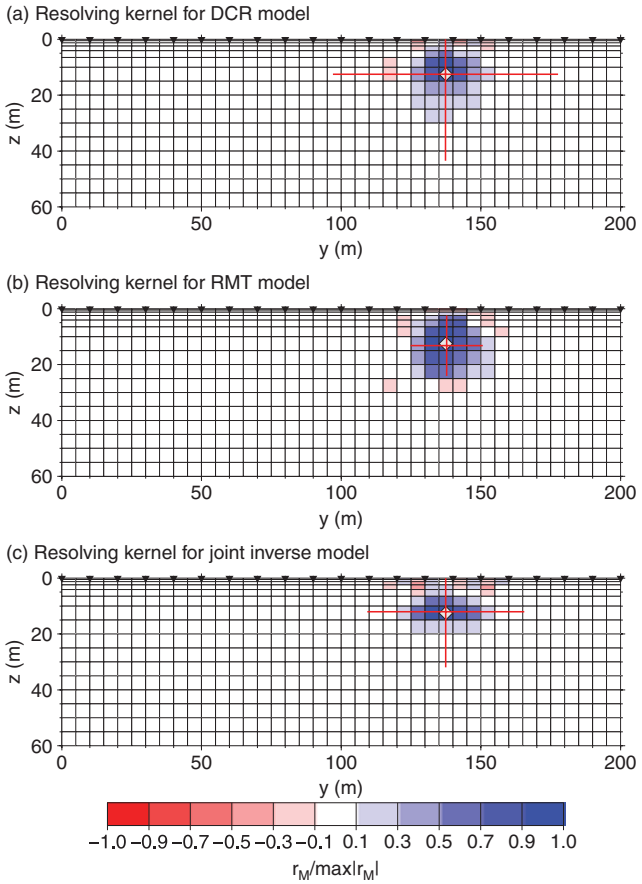

Figure 4. Resolving kernels derived from the smoothness-constrained scheme for cell C in the shallow resistor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. A white diamond marks the considered cell. The corresponding error estimates are given in Table 1. The resolving kernel and the most-squares errors of the joint inversion suggest that the resistivity of cell C is best constrained in the joint inversion.

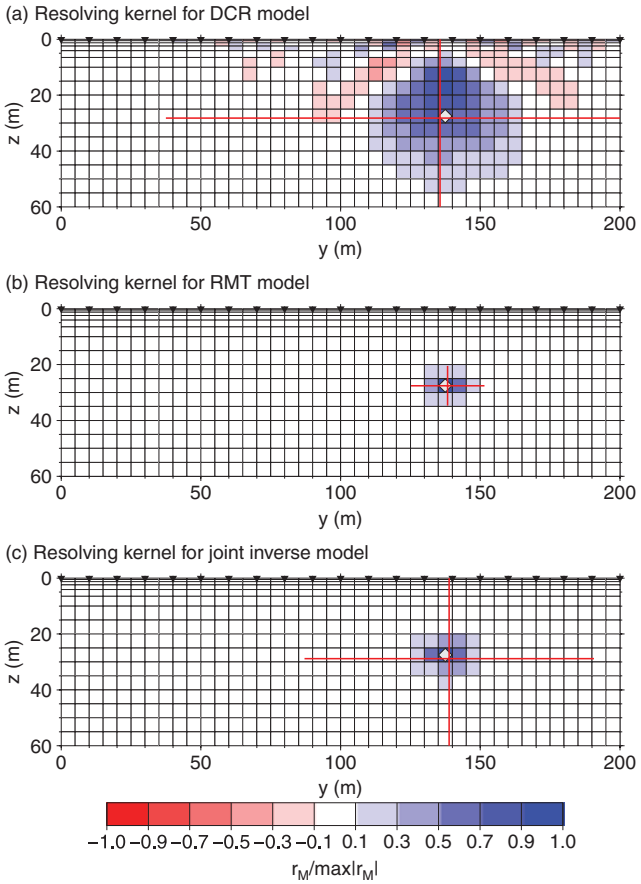

Figure 5. Resolving kernels derived from the smoothness-constrained scheme for cell D in the deep conductor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. A white diamond marks the considered cell. The corresponding error estimates are given in Table 1. Due to the limited length of the electrode array, cell D is not resolved by the DCR data. With the given data weighting, the resolving kernel and most-squares errors of the joint inverse model are close to those of the RMT inverse model.

significant non-linear effect for the smoothness constrained scheme. However, the smoothness constraints introduce a strong linear component, and the deviation of the linearized error estimates from the most-squares estimates is relatively small. In stochastic terms, the most-squares analysis only probes a limited range of models close to the maximum of the probability density function. As a consequence, the confidence level might deviate from 68 per cent with different severity for different model parameters, even though the linearized and most-squares errors are comparable. The resolving kernels determined by the smoothness-constrained approach (Figs 3–5) are in all cases quite spread even for conductive near-surface structures in the RMT model (Fig. 3b). For cell C, the joint inversion (Fig. 4c) improves model error and resolution in comparison to the DCR (Fig. 4a) and RMT (Fig. 4b) single inversions. In contrast to that, the resolution of cell D is mostly given through the RMT data. The resolving kernels of the RMT single inversion (Fig. 5b) and the joint inversion (Fig. 5c) are quite similar, whereas the resolving kernel of the DCR problem (Fig. 5a) exhibits a significantly larger spread.

For the TSVD scheme, analyses were done in subspaces with the shape of half-ellipses (*cf.* Kalscheuer & Pedersen 2007). The threshold f_{thresh} for errors derived from non-linear semi-axes was adjusted to obtain comparable most-squares errors in the single and joint inversions of $f^{\pm}_{MSQ} \approx 2$. The errors, as obtained from the non-linear semi-axes and the most-squares inversions, truncation levels and figure numbers of the corresponding resolving kernels are summarized in Table 2. The resolving kernels of cells A, C and D are depicted in Figs 7, 8 and 9, respectively.

A first insight into the non-linear nature of the DCR, RMT and joint inverse problems is obtained from a comparison of linear and non-linear semi-axes in Figs 6(a)–(c) for the subspace pertaining to cell C. In all inverse problems, the non-linear semi-axes follow the trend of the linear semi-axes for small singular value numbers, but gradually stop increasing before reaching a value of one. As for the RMT study by Kalscheuer & Pedersen (2007), this might endorse the expectation that non-linearity leads to DCR inverse models that are better constrained than expected from linear studies. For a given upper limit of the size of the semi-axes, the number of semi-axes that is below the limit is largest for the joint inversion. Hence, the joint inversion increases the number of important model parameters (Vozoff & Jupp 1975).

Table 2. Results of the model variance and resolution analyses with the TSVD scheme.

Data	Step of analysis	Cell A	Cell B	Cell C	Cell D
DCR	Resolving kernel	Fig. 7(a)	Fig. –	Fig. 8(a)	Fig. 9(a)
	min(M/N)	234	277	234	277
	Truncation level of final variance analysis	52	97	58	70
	Factors (f^-/f^+) of final variance analysis	1.29/1.29	1.30/1.30	1.29/1.27	1.28/1.27
	Factors (f^-_{MSQ}/f^+_{MSQ})	1.40/1.62	81.3/20.2	2.33/3.19	4.98/2.27
TE & TM	Resolving kernel	Fig. 7(b)	Fig.–	Fig. 8(b)	Fig. 9(b)
	min(M/N)	234	277	234	277
	Truncation level of final variance analysis	234	209	66	63
	Factors (f^-/f^+) of final variance analysis	1.75/1.76	1.99/2.00	1.28/1.27	1.30/1.29
	Factors (f^-_{MSQ}/f^+_{MSQ})	1.49/1.46	8.82/143.9	1.44/1.48	1.29/1.43
TE & TM & DCR	Resolving kernel	Fig. 7(c)	Fig.–	Fig. 8(c)	Fig. 9(c)
	min(M/N)	234	277	234	277
	Truncation level of final variance analysis	95	226	72	85
	Factors (f^-/f^+) of final variance analysis	1.38/1.40	1.86/1.86	1.26/1.26	1.27/1.27
	Factors (f^-_{MSQ}/f^+_{MSQ})	2.12/2.25	41.8/664.5	1.32/1.34	1.45/2.85

Notes: The threshold f_{thresh} for errors derived from non-linear semi-axes was adjusted to obtain comparable most-squares errors in the single and joint inversions of $f^{\pm}_{MSQ} \approx 2$. The resulting truncation levels, error factors $f^{\pm}$ from the non-linear semi-axes and error factors $f^{\pm}_{MSQ}$ from the most-squares inversion are given. For the DCR model, errors $f^{\pm}_{MSQ}$ are systematically larger than errors $f^{\pm}$. For the RMT model, errors $f^{\pm}_{MSQ}$ are in good agreement with errors $f^{\pm}$ for cells A and D pertaining to conductive structures within the depth of investigation range.

However, a comparison of errors derived from non-linear semi-axes and from most-squares inversions reveals important differences in the non-linear nature of the inverse problems:

(1) For the DCR method, error factors $f^{\pm}_{MSQ}$ derived from most-squares inversion are consistently larger than the errors $f^{\pm}$ computed from non-linear semi-axes (*cf.* Table 2). For a threshold value of $f_{\text{thresh}} = 2$ (not shown), the most-squares estimates were up to two orders of magnitude above the desired error level of two. Apparently, the description of the non-linear confidence surface only in the direction of the model eigenvectors gives misleading results of model variability. One reason for this discrepancy might be a significant degree of rotation of the model eigenvectors with position in parameter space. Also, the transformation to logarithmic model parameters does not appear to linearize the 2-D inverse problem as much as suggested by the 1-D examples in Johansen (1977). As there is no interest in models with unreasonably large errors, the threshold value was reduced to $f_{\text{thresh}} = 1.3$ (*cf.* Table 2). The relatively large spreads of the resolving kernels (Figs 8a and 9a) might be related to the lack of a vertical scale length in DCR models as illustrated by Parker (1984) with 1-D computations. The resolving kernels in Figs 7(a), 8(a) and 9(a) and the smooth inverse model in Fig. 2(b) suggest that the DCR inverse model is largely constrained by the imposed smoothness.

(2) For the RMT method, error factors $f^{\pm}$ from the non-linear semi-axes are verified by the most-squares errors $f^{\pm}_{MSQ}$ for most parameters within the depth range of investigation (*cf.* Table 2). Considering the resolving kernels in Figs 7–9, it can be concluded that especially conductive structures within the depth of investigation range are very well constrained by the data themselves.

(3) For the joint inverse problem, the most-squares error estimates $f^-_{MSQ}/f^+_{MSQ} = 1.32/1.34$ of cell C are smaller than the ones of the RMT single inversion with $f^-_{MSQ}/f^+_{MSQ} = 1.44/1.48$ and of the DCR single inversion where $f^-_{MSQ}/f^+_{MSQ} = 2.33/3.19$. At the same time, the resolving kernel of the joint inverse model (Fig. 8c) is less spread than the resolving kernels of the single inversions (Figs 8a and b). This means the resistivity of cell C is better constrained in the joint inversion than in the single inversion. The errors of cell D reflect the importance of a properly chosen data weighting. With the given data weighting, errors of $f^-_{MSQ}/f^+_{MSQ} = 1.45/2.85$ for the resistivity of cell D are obtained which are relatively close to $f^-_{MSQ}/f^+_{MSQ} = 1.29/1.43$ of the RMT model and significantly smaller than $f^-_{MSQ}/f^+_{MSQ} = 4.98/2.27$ of the DCR model. An equal weighting of all data points leads to increased errors $f^-_{MSQ}/f^+_{MSQ} = 2.65/3.89$ and a resolving kernel (Fig. 10) that has a larger spread.

It is instructive to consider most-squares models. Due to the coupling between model parameters, variations in the considered model parameter are partly compensated by correspondingly large variations of the other model parameters. For the single and joint inverse models, the TSVD-based most-squares models for minimal and maximal resistivities of cell C are shown in the left-hand and right-hand columns of Fig. 11, respectively. Obviously, the variation of model parameters is within reasonable limits for the given error levels of the resistivity of cell C. The variation of the resistivities of nearby cells increases with increasing most-squares error.

Regarding the data resolution matrix of the joint inverse problem, there is a clear pattern in individual rows (not shown) for TE-mode data, TM-mode data or DCR data:

(1) A row pertaining to a given TE-mode datum has large entries only for the considered datum and for other TE-mode data.

(2) In a row pertaining to a given TM-mode datum, there are large entries corresponding to the datum itself and to other TM-mode data at neighbouring stations and frequencies. Typically, entries pertaining to the TE-mode data are negligible. Some entries corresponding

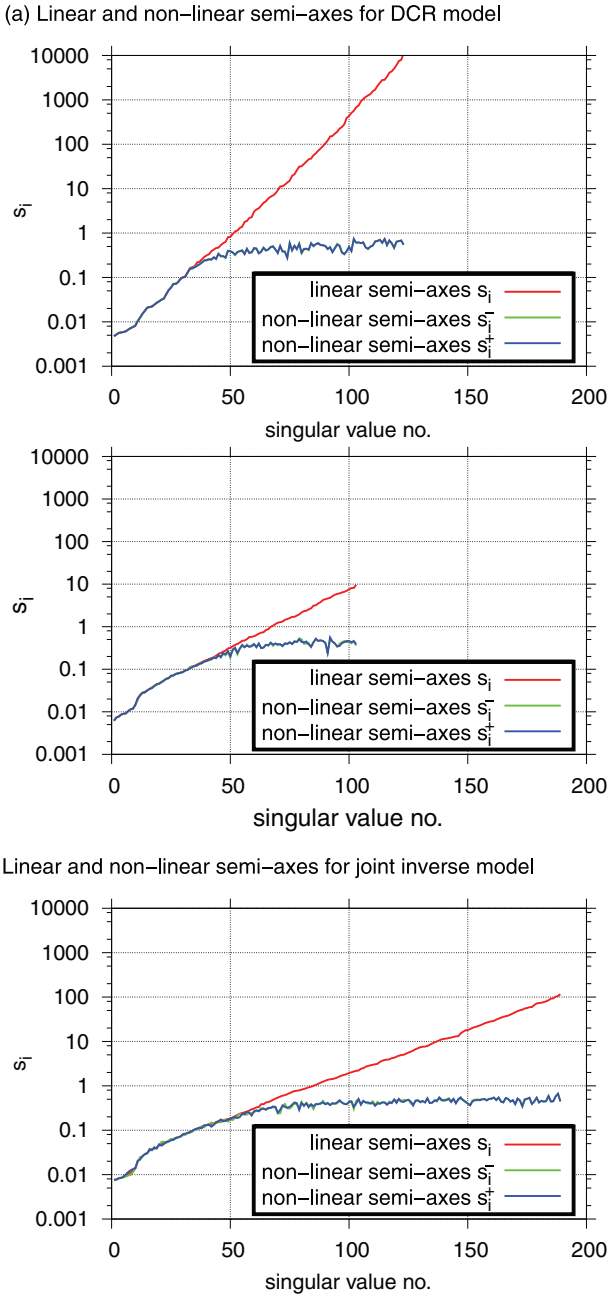

Figure 6. Linear and non-linear semi-axes of the single and joint inverse models for the subspace of cell C for (a) the DCR forward and reverse pole–dipole data, (b) the TE-mode and TM-mode data and (c) the combined data set. For illustrative purposes, the semi-axes for an error threshold $f_{\text{thresh}} = 2$ are shown. The number of semi-axes from the combined data set that is below a given threshold is significantly larger than for any single data set. Hence, in the joint inversion, the number of important model parameters is increased and the resistivity of cell C is better constrained.

to DCR data are non-negligible, though they are in amplitude not quite as large as those of the TM-mode data.

(3) A row pertaining to a given DCR datum displays strong coupling to other DCR data at neighbouring electrode positions and with similar electrode separation. The entries pertaining to TE-mode data are much smaller than those for TM-mode data.

These observations confirm the expectation (Monteiro Santos *et al.* 2007) that TM-mode and DCR data are compatible in the sense that the direction of current flow is mostly in the plane of the profile in both configurations and that both configurations are sensitive to resistive structures. For field data that are affected by 3-D distortions, however, it can be anticipated that the decoupling between TE-mode data on the one hand and TM-mode and DCR data on the other hand is less distinct.

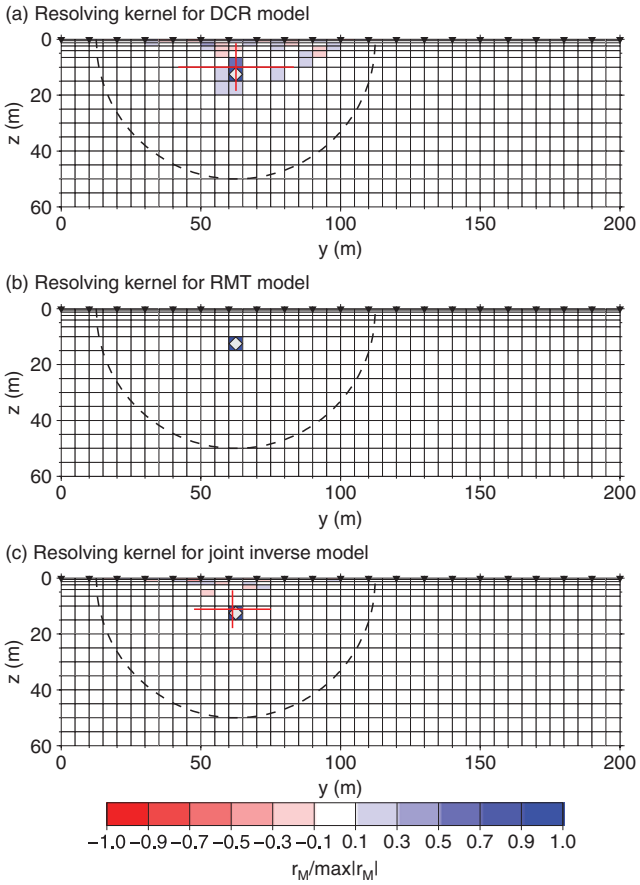

Figure 7. Resolving kernels derived from non-linear TSVD analyses for cell A in the shallow conductor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. The dashed line marks the border of the chosen subspace. A white diamond marks the considered cell. The corresponding truncation levels and non-linear error estimates are given in Table 2. The RMT data provide better resolving power for the conductor than the DCR data. For the joint inverse model, resolution and most-squares errors of parameter A are slightly degraded as compared to the RMT inversion.

4 DISCUSSION AND CONCLUSIONS

With two conceptually different analysis schemes, error and resolution properties are estimated for 2-D smoothness-constrained models derived from single and joint inversions of DCR and RMT data. Using a non-linear TSVD analysis and a TSVD-based most-squares inversion as well as a smoothness-constrained linearized analysis and a smoothness-constrained most-squares inversion, the effect of non-linearity on the model error and resolution estimates is investigated. The smoothness-constrained linearized analysis and most-squares inversion describe to which degree inverse models are determined by the combination of data and smoothness-constraints.

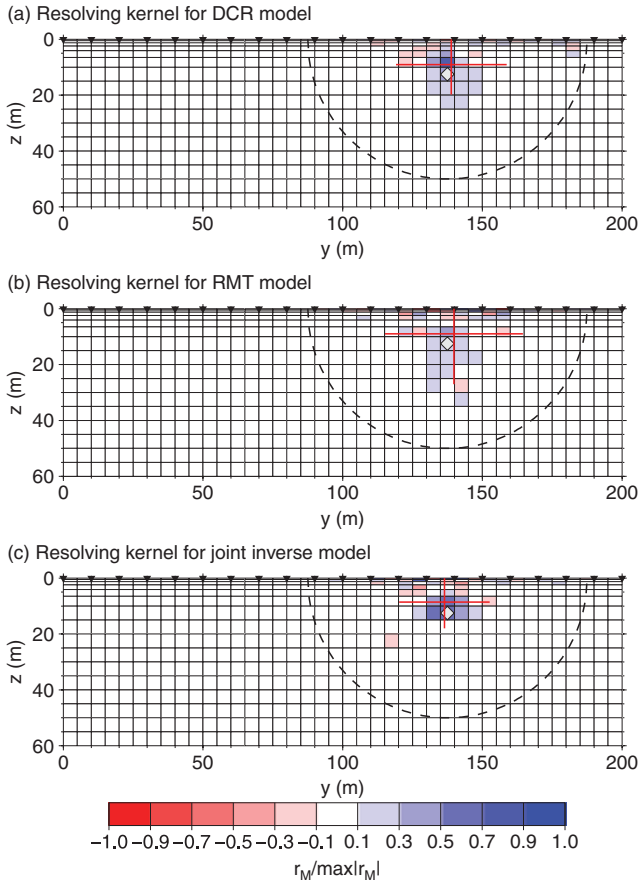

Figure 8. Resolving kernels derived from non-linear TSVD analyses for cell C in the shallow resistor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. The dashed line marks the border of the chosen subspace. A white diamond marks the considered cell. The corresponding truncation levels and non-linear error estimates are given in Table 2. The combined data set resolves the resistivity of cell C better than the single data sets with parameter errors that are smaller than those of the single inversions.

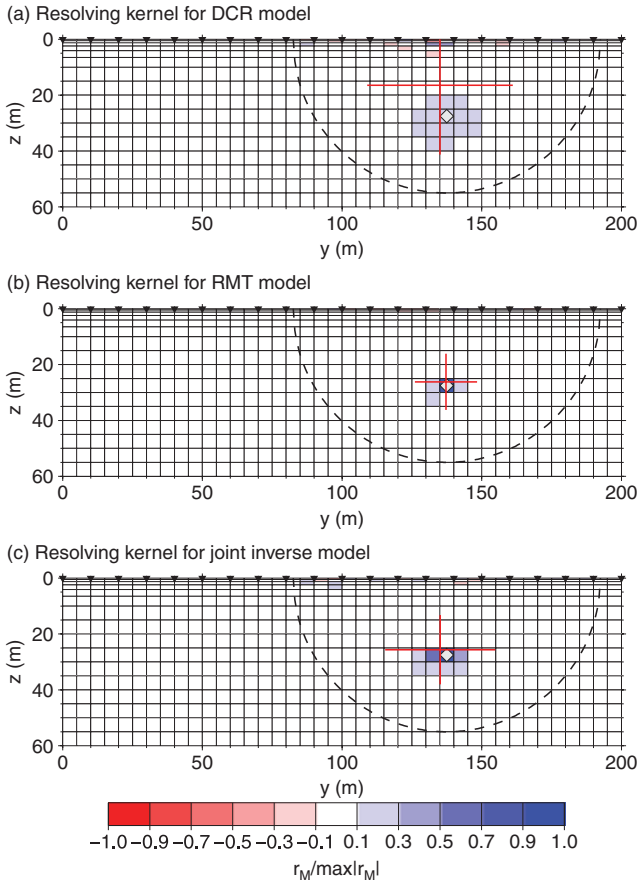

Figure 9. Resolving kernels derived from non-linear TSVD analyses for cell D in the deep conductor. The panels show the resolving kernels for the inverse models derived from (a) DCR forward and reverse pole–dipole data, (b) TE-mode and TM-mode data and (c) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. The dashed line marks the border of the chosen subspace. A white diamond marks the considered cell. The corresponding truncation levels and non-linear error estimates are given in Table 2. Cell D is outside the depth range of exploration for the chosen DCR configuration. The resolving kernel and error estimates of the joint inverse model are comparable to the ones of the RMT inverse model as the weight of the DCR data and high-frequency RMT data was reduced.

The TSVD-based schemes estimate to which degree inverse models can potentially be constrained by the data alone. In both schemes, model variability is investigated in a limited region in model parameter space around the model with minimum misfit. Outside this region the misfit function is not sampled and, hence, the confidence level associated with individual parameter errors might deviate from 68 per cent.

The smoothness-constrained most-squares inversion verifies the linearized smoothness-constrained parameter errors for most model parameters investigated here. This leads to the important conclusion that the widely used linearized smoothness-constrained parameter error estimates can be representative estimates of model variability in cases where model smoothness is a reasonable assumption. The reasonably good agreement between linearized and most-squares error estimates reflects the fact that the model regularization itself helps to linearize the inverse problem.

For DCR data, the TSVD-based most-squares inversion demonstrates that model variability is larger than estimated by the non-linear semi-axes. The consistent discrepancies between the different error estimates gives hints about the non-linear nature of the DCR inverse problem. First in the 2-D case, the transformation to

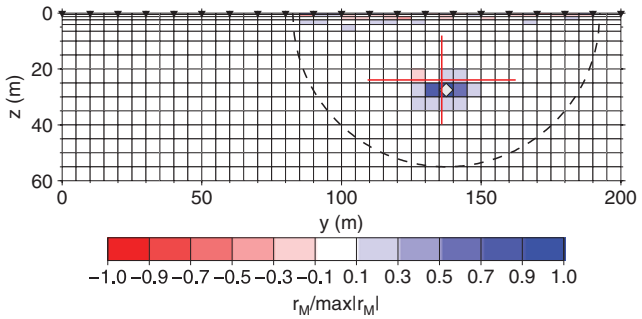

Figure 10. Resolving kernel derived from a non-linear TSVD analysis for cell D in the deep conductor for a joint inversion with equal weighting of all data. Clearly, the resolving kernel and the corresponding error estimates $f^{-}_{MSQ}/f^{+}_{MSQ} = 2.65/3.89$ demonstrate that equal weighting of DCR and RMT data leads to a resistivity of cell D that is less well constrained than in the joint inverse model with reweighted DCR data and high-frequency RMT data (*cf.* Table 2 and Fig. 9c).

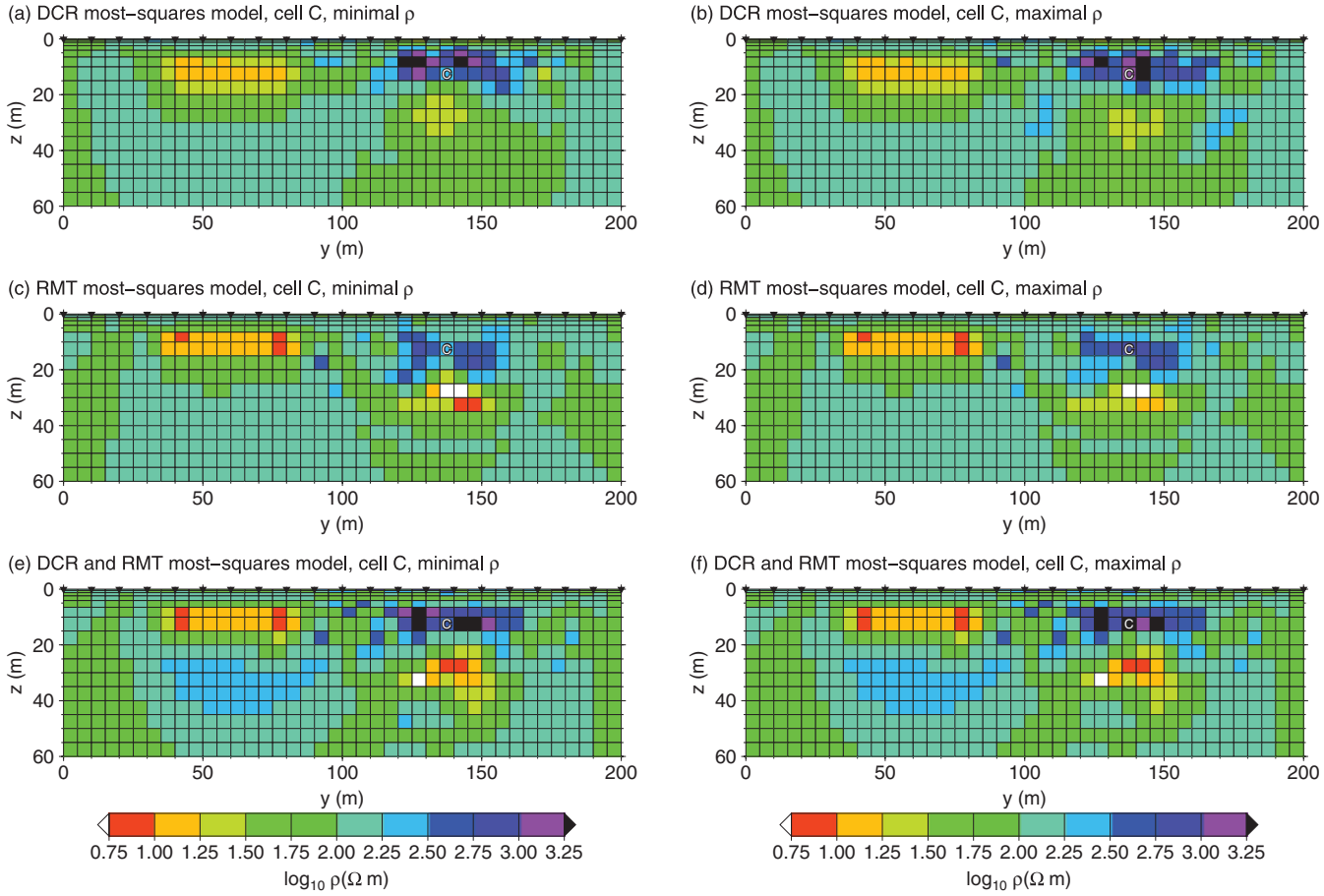

Figure 11. Most-squares models for the non-linear TSVD analyses of cell C in the shallow resistor. The extremal models for minimal and maximal resistivity of cell C are derived from (a) and (b) DCR forward and reverse pole–dipole data, (c) and (d) TE-mode and TM-mode data and (e) and (f) TE-mode, TM-mode and DCR forward and reverse pole–dipole data. The corresponding truncation levels and non-linear error estimates are given in Table 2. The most-squares models illustrate that the variation of model parameters is within reasonable limits for an error level of the resistivity of cell C that is approximately a factor two.

logarithmic model parameters does not appear to lead to such a large degree of linearization as described by Johansen (1977). Second, a significant degree of rotation of the model eigenvectors occurs. Third, at reasonable error levels, the relatively large spread of the resolving kernels appears to express the badly determined vertical scale length derived for 1-D DCR models by Parker (1984). It would be interesting to investigate to which degree model error and resolution could be improved by additionally employing borehole electrodes.

For RMT data, the description of the non-linear confidence surface with non-linear semi-axes yields model error estimates that are in good agreement with the errors form the most-squares inversion in an average sense. Hence, non-linearity causes inverse models to be better constrained than expected from studies based on linear semi-axes.

Qualitatively, the error and resolution estimates from the smoothness-constrained scheme and the non-linear TSVD scheme are in good agreement. For structures that are within the depth range of investigation of both RMT and DCR data (such as cell C in our synthetic example), both schemes show that a joint inversion can significantly improve model errors and resolution. The data sets need to be adequately weighted to ensure that error and resolution, as compared to the best single inversion result, are not overduely deteriorated in such parts of the model that are outside the depth of exploration range of one of the methods (such as cell D in our synthetic example).

ACKNOWLEDGMENTS

Weerachai Siripunvaraporn of Mahidol University, Thailand, and Gary Egbert of Oregon State University made the REBOCC code available for modifications. We greatly appreciate that we could use the Cholesky decomposition and SVD routines of the LAPACK library provided by Anderson *et al.* (1999). We would like to thank two anonymous reviewers and the editor Mark Everett for their comments which helped to improve the manuscript.

REFERENCES

Alumbaugh, D.L. & Newman, G.A., 2000. Image appraisal for 2-D and 3-D electromagnetic inversion, *Geophysics,* **65**(5), 1455–1467.

Anderson, E. *et al.*, 1999. *LAPACK Users' Guide,* 3rd edn, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Aprea, C., Booker, J.R. & Smith, J.T., 1997. The forward problem of electromagnetic induction: accurate finite-difference approximations for two-dimensional discrete boundaries with arbitrary geometry, *Geophys. J. Int.,* **129**(1), 29–40.

Candansayar, M.E. & Tezkan, B., 2008. Two-dimensional joint inversion of radiomagnetotelluric and direct current resistivity data, *Geophys. Prospect.,* **56**(5), 737–749.

Christensen, N.B., 2000. Difficulties in determining electrical anisotropy in subsurface investigations, *Geophys. Prospect.,* **48**(1), 1–19.

Christensen-Dalsgaard, J., Hansen, P.C. & Thompson, M.J., 1993. Generalized singular-value decomposition analysis of helioseismic inversions, *Mon. Not. R. astr. Soc.,* **264**(3), 541–564.

Constable, S.C., Parker, R.L. & Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics,* **52**(3), 289–300.

de Groot-Hedlin, C. & Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models for magnetotelluric data, *Geophysics,* **55**(12), 1613–1624.

Dey, A. & Morrison, H.F., 1979. Resistivity modelling for arbitrarily shaped two-dimensional structures, *Geophys. Prospect.,* **27**(1), 106–136.

Friedel, S., 2003. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach, *Geophys. J. Int.,* **153**(2), 305–316.

Golub, G. & van Loan, C., 1996. *Matrix Computations,* 3rd edn, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore and London.

Gómez-Trevino, E. & Edwards, R.N., 1983. Electromagnetic soundings in the sedimentary basin of southern Ontario: a case history, *Geophysics,* **48**(3), 311–330.

Haber, E. & Oldenburg, D., 1997. Joint inversion: a structural approach, *Inverse Probl.,* **13**(1), 63–77.

Hohmann, G.W., 1987. Numerical modeling for electromagnetic methods in geophysics, in *Electromagnetic Methods in Applied Geophysics,* Vol. 1, Theory, pp. 313–363, ed. Nabighian, M.N., SEG, Tulsa.

Hohmann, G.W. & Raiche, A.P., 1987. Inversion of controlled-source electromagnetic data, in *Electromagnetic Methods in Applied Geophysics,* Vol. 1, Theory, pp. 469–503, ed. Nabighian, M.N., SEG, Tulsa.

Jackson, D.D., 1973. Marginal solutions to quasilinear inverse problems in geophysics: the edgehog method, *Geophys. J. R. astr. Soc.,* **35**(1–3), 121–136.

Jackson, D.D., 1976. Most squares inversion, *J. geophys. Res.,* **81**(5), 1027–1030.

Johansen, H.K., 1977. A man/computer interpretation system for resistivity soundings over a horizontally stratified earth, *Geophys. Prospect.,* **25**(4), 667–691.

Jupp, D.L.B. & Vozoff, K., 1975. Simple iterative methods for the inversion of geophysical data, *Geophys. J. R. astr. Soc.,* **42**(3), 957–976.

Kalscheuer, T. & Pedersen, L.B., 2007. A non-linear truncated SVD variance and resolution analysis of two-dimensional magnetotelluric models, *Geophys. J. Int.,* **169**(2), 435–447.

Kalscheuer, T., Pedersen, L.B. & Siripunvaraporn, W., 2008. Radiomagnetotelluric two-dimensional forward and inverse modelling accounting for displacement currents, *Geophys. J. Int.,* **175**(2), 486–514.

LaBrecque, D.J., Morelli, G., Daily, W., Ramirez, A. & Lundegard, P., 1999. Occam's inversion of 3D ERT data, in *Three-Dimensional Electromagnetics,* pp. 575–590, ed. Spies, B., SEG, Tulsa.

Lines, L.R. & Treitel, S., 1984. Tutorial: a review of least-squares inversion and its application to geophysical problems, *Geophys. Prospect.,* **32**(2), 159–186.

Lines, L.R. & Treitel, S., 1985. Inversion with a grain of salt, *Geophysics,* **50**(1), 99–109.

Meju, M.A., 1994. Biased estimation: a simple framework for inversion and uncertainty analysis with prior information, *Geophys. J. Int.,* **119**(2), 521–528.

Meju, M.A., 2009. Regularized extremal bounds analysis (REBA): an approach to quantifying uncertainty in nonlinear geophysical inverse problems, *Geophys. Res. Lett.,* **36**(3), L03304.

Meju, M.A. & Hutton, V.R.S., 1992. Iterative most-squares inversion: application to magnetotelluric data, *Geophys. J. Int.,* **108**(3), 758–766.

Menke, W., 1989. Geophysical data analysis: discrete inverse theory, Vol. 45 of *International Geophysics Series*, Academic Press, London.

Miller, C.R. & Routh, P.S., 2007. Resolution analysis of geophysical images: comparison between point spread function and region of data influence measures, *Geophys. Prospect.,* **55**(6), 835–852.

Monteiro Santos, F.A., Afonso, A.R.A. & Dupis, A., 2007. 2D joint inversion of dc and scalar audio-magnetotelluric data in the evaluation of low enthalpy geothermal fields, *J. geophys. Eng.,* **4,** 53–62.

Muñoz, G. & Rath, V., 2006. Beyond smooth inversion: the use of nullspace projection in MT, *Geophys. J. Int.,* **164**(2), 301–311.

Oldenburg, D.W. & Li, Y.G., 1999. Estimating depth of investigation in dc resistivity and IP surveys, *Geophysics,* **64**(2), 403–416.

Ory, J. & Pratt, R.G., 1995. Are our parameter estimators biased? The significance of finite-difference regularization operators, *Inverse Probl.,* **11**(2), 397–424.

Parker, R.L., 1977. Understanding inverse theory, *Ann. Rev. Earth Planet. Sci.,* **5,** 35–64.

Parker, R.L., 1984. The inverse problem of resistivity sounding, *Geophysics,* **49**(12), 2143–2158.

Pedersen, L.B., 1979. Constrained inversion of potential field data, *Geophys. Prospect.,* **27**(4), 726–748.

Pedersen, L.B. & Rasmussen, T.M., 1989. Inversion of magnetotelluric data: a non-linear least squares approach, *Geophys. Prospect.,* **37**(6), 669–695.

Raiche, A.P., Jupp, D.L.B., Rutter, H. & Vozoff, K., 1985. The joint use of coincident loop transient electromagnetic and Schlumberger sounding to resolve layered structures, *Geophysics,* **50**(10), 1618–1627.

Rodi, W. & Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion, *Geophysics,* **66**(1), 174–187.

Rodi, W.L., 1976. A technique for improving the accuracy of finite element solutions for magnetotelluric data, *Geophys. J. R. astr. Soc.,* **44**(2), 483–506.

Sasaki, Y., 1989. Two-dimensional joint inversion of magnetotelluric and dipole-dipole resistivity data, *Geophysics,* **54**(2), 254–262.

Schwalenberg, K., Rath, V. & Haak, V., 2002. Sensitivity studies applied to a two-dimensional resistivity model from the Central Andes, *Geophys. J. Int.,* **150**(3), 673–686.

Siripunvaraporn, W. & Egbert, G., 2000. An efficient data-subspace inversion method for 2-D magnetotelluric data, *Geophysics,* **65**(3), 791–803.

Vozoff, K. & Jupp, D.L.B., 1975. Joint inversion of geophysical data, *Geophys. J. R. astr. Soc.,* **42**(3), 977–991.

Weaver, J.T., Le Quang, B.V. & Fischer, G., 1985. A comparison of analytical and numerical results for a two-dimensional control model in electromagnetic induction - I: B-polarization calculations, *Geophys. J. R. astr. Soc.,* **82**(2), 263–277.

Weaver, J.T., Le Quang, B.V. & Fischer, G., 1986. A comparison of analytical and numerical results for a two-dimensional control model in electromagnetic induction - II: E-polarization calculations, *Geophys. J. R. astr. Soc.,* **87**(3), 917–948.

Xu, S., Duan, B. & Zhang, D., 2000. Selection of the wavenumbers k using an optimization method for the inverse Fourier transform in 2.5D electrical modelling, *Geophys. Prospect.,* **48**(5), 789–796.

APPENDIX A: INVERSE THEORY

In an inverse problem, a vector of N forward responses $\mathbf{F}[\mathbf{m}] = (F_1, \ldots, F_N)^{\mathrm{T}}$ is to be fitted to a vector of N measurements $\mathbf{d} = (d_1, \ldots, d_N)^{\mathrm{T}}$ with corresponding standard deviations $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N)^{\mathrm{T}}$ by modifying a vector of M model parameters $\mathbf{m} = (m_1, \ldots, m_M)^T$ (Jackson 1973; Parker 1977). To reduce the non-uniqueness and ill-posedness of the inverse problem (Parker 1977), one minimizes a combination of a data misfit measure Q_{d} and a solution simplicity measure Q_m in the form of an unconstrained functional (Menke 1989)

$$U[\mathbf{m}, \lambda] = \left(Q_{\mathrm{d}}[\mathbf{m}] - Q_{\mathrm{d}}^{*}\right) + \lambda Q_m[\mathbf{m}], \tag{A1}$$

where

$$Q_{\mathrm{d}}[\mathbf{m}] = (\mathbf{d} - \mathbf{F}[\mathbf{m}])^{\mathrm{T}} \mathbf{W}_{\mathrm{d}}^{\mathrm{T}} \mathbf{W}_{\mathrm{d}} (\mathbf{d} - \mathbf{F}[\mathbf{m}]), \tag{A2}$$

$$Q_m[\mathbf{m}] = (\mathbf{m} - \mathbf{m}_r)^{\mathrm{T}} \mathbf{W}_m^{\mathrm{T}} \mathbf{W}_m (\mathbf{m} - \mathbf{m}_r). \tag{A3}$$

The superscript T denotes a matrix transpose, $\mathbf{W}_\mathrm{d} = \mathrm{diag}\{\sigma_1^{-1}, \ldots, \sigma_N^{-1}\}$ is a data weighting matrix, and Q_d^* is the target data misfit. The forward operator $\mathbf{F}[\mathbf{m}]$ is generally non-linearly related to $\mathbf{m}$. The matrix $\mathbf{W}_m$ of model regularizing operators describes the pre-conceived simplicity of the solution $\mathbf{m}$ relative to a reference model $\mathbf{m}_r$. The Lagrange multiplier λ describes the balance between data fit ($Q_\mathrm{d}[\mathbf{m}]$) and solution simplicity ($Q_m[\mathbf{m}]$).

By means of a linearization of the forward operator in the vicinity of the current model $\mathbf{m}_k$, the original problem of minimizing the quantity U is replaced through that of minimizing a functional U^lin which is quadratic in $\mathbf{m}_{k+1}$ (Lines & Treitel 1984; Menke 1989):

$$U^\mathrm{lin}[\mathbf{m}_{k+1}, \lambda] = \left(Q_d^\mathrm{lin}[\mathbf{m}_{k+1}] - Q_d^*\right) + \lambda Q_m[\mathbf{m}_{k+1}], \tag{A4}$$

where

$$\begin{aligned} Q_\mathrm{d}^\mathrm{lin}[\mathbf{m}_{k+1}] &= (\mathbf{d} - \mathbf{F}[\mathbf{m}_k] - \mathbf{J}(\mathbf{m}_{k+1} - \mathbf{m}_k))^\mathrm{T} \mathbf{W}_\mathrm{d}^\mathrm{T} \\ &\quad \mathbf{W}_\mathrm{d}(\mathbf{d} - \mathbf{F}[\mathbf{m}_k] - \mathbf{J}(\mathbf{m}_{k+1} - \mathbf{m}_k)), \qquad \text{(A5)} \\ \mathbf{J} &= \{\partial F_i[\mathbf{m}]/\partial m_j\}_{\mathbf{m}=\mathbf{m}_k}, \end{aligned}$$

$\mathbf{J}$ is the Jacobian matrix of partial derivatives at the model $\mathbf{m}_k$ of the kth iteration.

Setting the gradient $\nabla_{\mathbf{m}_{k+1}} U^\mathrm{lin}[\mathbf{m}_{k+1}, \lambda] = 0$ yields the approximate solution in eq. (1).

A1 Smoothness-constrained inversion

In a smoothness-constrained inversion (Constable *et al.* 1987; de Groot-Hedlin & Constable 1990; Ory & Pratt 1995), the matrix $\mathbf{W}_m^\mathrm{T}\mathbf{W}_m$ is replaced by a 2-D flattening matrix $\mathbf{S} = \alpha_y \partial_y^\mathrm{T}\partial_y + \alpha_z \partial_z^\mathrm{T}\partial_z$ where ∂_y and ∂_z are horizontal and vertical gradient matrices, respectively, and α_y and α_z are factors that permit different weighting of horizontal and vertical smoothness, respectively (de Groot-Hedlin & Constable 1990). Often, there is no assumption about a reference model such that $\mathbf{m}_r = \mathbf{0}$ and the flatness of the estimated model is the essence of model regularization.

A2 TSVD inversion

In the Marquardt-Levenberg method (Lines & Treitel 1984), the inverse problem is regularized by requiring the minimum solution length. Hence, the reference model is that of the previous iteration $\mathbf{m}_r = \mathbf{m}_k$, the model regularizing matrix is an identity matrix $\mathbf{W}_m = \mathbf{I}$, and the Lagrange multiplier λ is an iteration dependent damping factor ϵ_{k+1}.

Typically, the generalized inverse of the Marquardt-Levenberg method is constructed in terms of a singular value decomposition (SVD) of the normalized Jacobian $\mathbf{J}_\mathbf{W} = \mathbf{W}_\mathrm{d}\mathbf{J}$ (Jupp & Vozoff 1975; Lines & Treitel 1984; Menke 1989; Golub & van Loan 1996). To further stabilize the solution, only the $p \leq \min(M, N)$ largest singular values might be included in the solution process. Hence, the generalized inverse becomes

$$\begin{aligned} \mathbf{J}_\mathbf{W}^{-\mathrm{g}} &= \mathbf{V}_p \tilde{\mathbf{\Lambda}}_p^{-1} \mathbf{U}_p^\mathrm{T} \\ &= \mathbf{V}_p \mathrm{diag}\left(\frac{\lambda_i}{\lambda_i^2 + \epsilon_{k+1}}\right)_{i=1,\ldots,p} \mathbf{U}_p^\mathrm{T}, \end{aligned} \tag{A6}$$

where the $\lambda_i \in \mathbb{R}, i = 1, \ldots, p$ are the p largest singular values and the columns of the matrices $\mathbf{U}_p \in \mathbb{R}^{N*p}$ and $\mathbf{V}_p \in \mathbb{R}^{M*p}$ are the data and model eigenvectors pertaining to the p largest singular values.

Solutions produced by the damped TSVD scheme are strongly influenced by the initial model and oscillations with large amplitudes may arise in overdiscretized models (Constable *et al.* 1987; Christensen-Dalsgaard *et al.* 1993; Ory & Pratt 1995). Hence, an inversion with the Occam scheme is preferred to a damped TSVD inversion for 2-D inverse problems with arbitrarily chosen starting models.

APPENDIX B: COMPUTATION OF FORWARD RESPONSES AND SENSITIVITY MATRICES

B1 DCR case

Forward responses are computed with the finite-difference approximation (FDA) by Dey & Morrison (1979), that yields a discretization by area. As the DCR source field is 3-D, a Fourier transform along the strike direction is performed that renders the original 3-D forward problem a 2.5-D forward problem. The wavenumbers involved in the Fourier-transformation are computed with the optimization scheme by Xu *et al.* (2000) which operates with a reduced set of wavenumbers. In this study five wavenumbers are used. The sensitivities of DCR data with respect to the model parameters are computed with the method by LaBrecque *et al.* (1999) which utilizes the reciprocity of the forward problem, that is, the symmetry of the system matrix of the forward problem, to reduce computing time.

B2 RMT case

The forward responses are computed with a finite-difference approximation of the Helmholtz equations for the transverse electric (TE) and transverse magnetic (TM) modes (Hohmann 1987; Aprea *et al.* 1997). The solution of the resulting systems of linear equations yields the electric and magnetic field components along the strike direction, that is, E_x and H_x. The remaining auxiliary field components that are necessary to compute the TE-mode and TM-mode impedances and the vertical magnetic transfer function are obtained from Taylor expansions of E_x and H_x in both vertical and horizontal directions up to second order and by substituting the appropriate Helmholtz equations for the second-order terms (Weaver *et al.* 1985, 1986). The sensitivity matrix is computed with the sensitivity-equation approach by Rodi (1976) and Rodi & Mackie (2001) that, similar to the DCR scheme by LaBrecque *et al.* (1999), accelerates computations by taking advantage of the reciprocity property of the forward problem. An elaborate description of the computation of RMT FDA forward responses and sensitivity matrices that allows for displacement currents can be found in Kalscheuer *et al.* (2008).

B3 Data alignment and weighting in joint inverse schemes

Designating the field data vectors, forward data vectors and Jacobian matrices of the RMT and DCR methods as $\mathbf{d}_\mathrm{RMT}$, $\mathbf{F}_\mathrm{RMT}[\mathbf{m}]$, $\mathbf{J}_\mathrm{RMT}$ and $\mathbf{d}_\mathrm{DCR}$, $\mathbf{F}_\mathrm{DCR}[\mathbf{m}]$ and $\mathbf{J}_\mathrm{DCR}$, respectively, the combined data vectors and Jacobian matrices that enter into eqs (A2), (A4) and (1) are given as

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}_\mathrm{RMT} \\ \mathbf{d}_\mathrm{DCR} \end{bmatrix}, \quad \mathbf{F}[\mathbf{m}] = \begin{bmatrix} \mathbf{F}_\mathrm{RMT}[\mathbf{m}] \\ \mathbf{F}_\mathrm{DCR}[\mathbf{m}] \end{bmatrix} \quad \text{and} \quad \mathbf{J} = \begin{bmatrix} \mathbf{J}_\mathrm{RMT} \\ \mathbf{J}_\mathrm{DCR} \end{bmatrix}. \tag{B1}$$

The inverse model obtained from a joint inversion might strongly be biased towards the single inversion model obtained from one of the methods. In the likely case that the number of DCR data is larger than the number of RMT data, the joint inverse model may closely resemble the DCR inverse model. Similarly, a highly resistive subsurface might bias the joint inverse model towards that of the DCR inverse model, as the DCR method is more sensitive to resistive structures than the RMT method. The influence of conductive structures embedded in the resistive host that is primarily seen in the RMT data might then be suppressed by the chosen model regularization due to the decreased relative sensitivity for the conductor.

To reduce the bias in joint inverse models, the importance of specific data can be modified through a diagonal data reweighting matrix $\mathbf{W}_{\text{rew}}$ (Candansayar & Tezkan 2008). The effective data weighting matrix is then given as

$$\mathbf{W}_{\text{d}} = \mathbf{W}_{\text{rew}} \begin{bmatrix} \mathbf{W}_{\text{d,RMT}} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{\text{d,DCR}} \end{bmatrix}, \tag{B2}$$

where $\mathbf{W}_{\text{d,RMT}}$ and $\mathbf{W}_{\text{d,DCR}}$ are weighting matrices of the RMT and DCR data.

If the data errors contained in the matrices $\mathbf{W}_{\text{d,RMT}}$ and $\mathbf{W}_{\text{d,DCR}}$ are close to the true ones, there is a certain danger of obtaining either a badly resolved or an unstable model when iterating to an rms of one for the reweighted system. Essentially, the desired rms should be adopted to the chosen weighting factors. Given that N_c of a total of N data are reweighted by a constant factor w_c and that the errors of $N_u = N - N_c$ data remain unchanged the rms takes the following form [*cf.* eq. (A2)]:

$$\begin{aligned} &\text{rms} \\ &= \sqrt{\frac{1}{N_u + N_c}\left\{\sum_{i=1}^{N_u}\left(\frac{d_i - F_i[\mathbf{m}]}{\sigma_i}\right)^2 + \sum_{i=N_u+1}^{N_u+N_c}\left(\frac{d_i - F_i[\mathbf{m}]}{w_c\sigma_i}\right)^2\right\}} \\ &= \sqrt{\frac{1}{N_u + N_c}\left\{Q_{du}^1 + Q_{dc}\right\}} \\ &= \sqrt{\frac{1}{N_u + N_c}\left\{Q_{du}^1 + \frac{1}{w_c^2}Q_{du}^2\right\}}. \end{aligned} \tag{B3}$$

If the data errors σ_i are correctly estimated, a reasonable fit to the data is obtained for $Q_{du}^1 \gtrapprox N_u$ and $Q_{du}^2 \gtrapprox N_c$. Hence, the desired rms should be chosen as

$$\text{rms}^{\text{opt}} \gtrapprox \sqrt{\frac{1}{N_u + N_c}\left\{N_u + \frac{1}{w_c^2}N_c\right\}}. \tag{B4}$$

APPENDIX C: DERIVATION OF MOST-SQUARES INVERSE SCHEME

To generate iterative model updates coming successively closer to the desired misfit threshold, Marquardt-Levenberg damping is used as a stabilizer in the TSVD-based and the smoothness-constrained most-squares schemes. This modification can be considered an extension to the smoothness-constrained most-squares inversion presented in Meju & Hutton (1992) and Meju (1994). The cost functional of a stabilized iterative most-squares scheme is then written as

$$\begin{aligned} U[\mathbf{m}_{k+1}, \mu, \epsilon_{k+1}] &= \mathbf{m}_{k+1}^{\text{T}}\hat{\mathbf{n}} + \\ &\quad \frac{1}{2\mu}\left\{\left(Q_{\text{d}}^{\text{lin}}[\mathbf{m}_{k+1}] + \lambda \cdot Q_m[\mathbf{m}_{k+1}] - Q_t\right)\right. \\ &\quad \left. + \epsilon_{k+1} \cdot Q_{ML}[\mathbf{m}_{k+1}]\right\} \\ Q_{\text{d}}^{\text{lin}}[\mathbf{m}_{k+1}] &= (\mathbf{d} - \mathbf{F}[\mathbf{m}_k] - \mathbf{J}(\mathbf{m}_{k+1} - \mathbf{m}_k))^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}} \\ &\quad \mathbf{W}_{\text{d}}(\mathbf{d} - \mathbf{F}[\mathbf{m}_k] - \mathbf{J}(\mathbf{m}_{k+1} - \mathbf{m}_k)), \\ Q_m[\mathbf{m}_{k+1}] &= (\mathbf{m}_{k+1} - \mathbf{m}_r)^{\text{T}}\mathbf{W}_m^{\text{T}}\mathbf{W}_m(\mathbf{m}_{k+1} - \mathbf{m}_r) \\ Q_{ML}[\mathbf{m}_{k+1}] &= (\mathbf{m}_{k+1} - \mathbf{m}_k)^{\text{T}}(\mathbf{m}_{k+1} - \mathbf{m}_k). \end{aligned} \tag{C1}$$

The Lagrange multiplier $\frac{1}{2\mu}$ describes the weighting between the parameter to be maximized or minimized and the misfit function $Q[\mathbf{m}_{k+1}] = Q_{\text{d}}^{\text{lin}}[\mathbf{m}_{k+1}] + \lambda \cdot Q_m[\mathbf{m}_{k+1}]$ as well as the Marquardt–Levenberg term $\epsilon_{k+1} \cdot Q_{ML}[\mathbf{m}_{k+1}]$ with a damping factor ϵ_{k+1}. The quantities λ, $\mathbf{W}_m$ and $\mathbf{m}_r$ appear only in the smoothness-constrained scheme.

The minimum of the cost functional in eq. (C1) is found to correspond to a model

$$\begin{aligned} \mathbf{m}_{k+1}(\mu, \lambda) &= \left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1} \\ &\quad \left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\hat{\mathbf{d}}_k - \mu\hat{\mathbf{n}} + \epsilon_{k+1}(\mathbf{m}_k - \mathbf{m}_r)\right) + \mathbf{m}_r \\ &= \mathbf{m}_{k+1}^{ls} - \left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\mu\hat{\mathbf{n}}, \end{aligned} \tag{C2}$$

which consists of the least-squares term $\mathbf{m}_{k+1}^{ls}$ [where $\hat{\mathbf{d}}_k = \mathbf{d} - \mathbf{F}[\mathbf{m}_k] + \mathbf{J}(\mathbf{m}_k - \mathbf{m}_r)$] and a compensating term related to the parameter to be maximized or minimized.

The Lagrange multiplier μ is determined by the quadratic constraint that the total linearized misfit

$$Q^{\text{lin}}[\mathbf{m}_{k+1}, \lambda] = Q_{\text{d}}^{\text{lin}}[\mathbf{m}_{k+1}] + \lambda \cdot Q_m[\mathbf{m}_{k+1}] + \epsilon_{k+1} \cdot Q_{ML}[\mathbf{m}_{k+1}] \tag{C3}$$

equals the threshold misfit Q_t. Substitution of eq. (C2) into eq. (C3) yields similar to Meju & Hutton (1992) and Meju (1994)

$$\mu = \pm\sqrt{\frac{Q_t - Q_{ls}^{\text{lin}}}{\hat{\mathbf{n}}^{\text{T}}\left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}}}}, \tag{C4}$$

where $Q_{ls}^{\text{lin}} = Q^{\text{lin}}[\mathbf{m}_{k+1}^{ls}, \lambda]$ is the linearized misfit function pertaining to the least-squares term $\mathbf{m}_{k+1}^{ls}$ from eq. (C2).

In the smoothness-constrained scheme, solely fixing the Lagrange multiplier λ and avoiding Marquardt-Levenberg damping does not grant the stability of the most-squares inversion, since the obtained misfits are often larger than the threshold value Q_t. The relatively small desired misfit deviation $\Delta Q = 1$ suggests small model variations at least for the well-determined model parameters. The additional Marquardt–Levenberg damping limits the model updates to the vicinity of the model of the previous iteration.

According to eqs (C2) and (C4), the model deviation away from the least-squares solution $\mathbf{m}_{k+1}^{ls}$ is

$$\begin{aligned} \Delta\mathbf{m}_{k+1} &= -\left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\mu\hat{\mathbf{n}} \\ &= \mp\sqrt{\frac{Q_t - Q_{ls}^{\text{lin}}}{\hat{\mathbf{n}}^{\text{T}}\left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}}}} \\ &\quad \cdot \left(\mathbf{J}^{\text{T}}\mathbf{W}_{\text{d}}^{\text{T}}\mathbf{W}_{\text{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\text{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}}. \end{aligned} \tag{C5}$$

The *j*th component of this deviation is

$$\begin{aligned}\hat{\mathbf{n}}^{\mathrm{T}}\Delta\mathbf{m}_{k+1} &= \mp\sqrt{\frac{Q_t - Q_{ls}^{\mathrm{lin}}}{\hat{\mathbf{n}}^{\mathrm{T}}\left(\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}}}} \\ &\quad \cdot \hat{\mathbf{n}}^{\mathrm{T}}\left(\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}} \\ &= \mp\sqrt{Q_t - Q_{ls}^{\mathrm{lin}}} \\ &\quad \sqrt{\hat{\mathbf{n}}^{\mathrm{T}}\left(\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m + \epsilon_{k+1}\mathbf{I}\right)^{-1}\hat{\mathbf{n}}}. \end{aligned} \tag{C6}$$

Cumulatively for all most-squares iterations, the deviation from the investigated model is the non-linear equivalent to $\sqrt{\hat{\mathbf{n}}^{\mathrm{T}}(\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m)^{-1}\hat{\mathbf{n}}}$. Assuming $[\mathrm{cov}\,\mathbf{m}_r] = (\lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m)^{-1}$, eq. (4) can be manipulated into the form $[\mathrm{cov}\,\mathbf{m}_{k+1}] = (\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}^{\mathrm{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J} + \lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m)^{-1}$. Hence, the former expression is equivalent to the standard deviation in form of the square root of the (j, j)th element of eq. (4). As the most-squares solution is constructed irrespective of the existence of the inverse of $\lambda\mathbf{W}_m^{\mathrm{T}}\mathbf{W}_m$, comparisons of most-squares results to linearized errors must be based on both terms of eq. (4).

The damping factor ϵ_{k+1} must be chosen carefully since it influences the direction of the model correction step. Although the damping factor can be chosen such that the confidence surface corresponding to $Q[\mathbf{m}^{\mathrm{opt}}] + \Delta Q$ is reached with one iteration, there is no guarantee that the resistivity of the considered cell is maximized or minimized by doing so. Therefore, an iterative implementation of the most-squares algorithm has to be applied. We limit the range of permissible damping factors to those, which result in misfits that do not exceed $Q[\mathbf{m}^{\mathrm{opt}}] + \Delta Q$ and are not smaller than the misfit of the previous iteration plus a small misfit increase, that enforces convergence. Reasonable choices are damping factors which give the largest ratio in resistivity change of the considered parameter to average resistivity change of the model or, more conservatively, the largest permissible damping factor or the smallest permissible misfit.